ON EIGENVALUES OF RECTANGULAR MATRICES

JULIUS BORCEA, BORIS SHAPIRO, AND MICHAEL SHAPIRO

ABSTRACT. Given a $(k+1)$-tuple $A, B_1, \ldots, B_k$ of $(m \times n)$-matrices with $m \leq n$ we call the set of all k-tuples of complex numbers $\{\lambda_1, \ldots, \lambda_k\}$ such that the linear combination $A + \lambda_1 B_1 + \lambda_2 B_2 + \ldots + \lambda_k B_k$ has rank smaller than m the *eigenvalue locus* of the latter pencil. Motivated primarily by applications to multi-parameter generalizations of the Heine-Stieltjes spectral problem, see [He] and [Vol], we study a number of properties of the eigenvalue locus in the most important case $k = n - m + 1$.

INTRODUCTION AND MAIN RESULTS

In recent years there appeared a number of publications discussing the eigenvalues of pencils of non-square matrices and their approximations, see, e.g., [BEGM], [CG], [TW] and references therein. But to the best of our knowledge the following natural problem either has been overlooked by specialists in linear algebra or is deeply buried in the (enormous) literature on this topic.

Question. Given a $(k+1)$-tuple of $(m \times n)$-matrices $A, B_1, \ldots, B_k$, $m \leq n$, describe the set of all values of the parameters $\lambda_1, \ldots \lambda_k$ for which the rank of the linear combination $A + \lambda_1 B_1 + \ldots + \lambda_k B_k$ is less than m or, in other words, when the linear system $v * (A + \lambda_1 B_1 + \ldots + \lambda_k B_k) = 0$ has a nontrivial left solution $0 \neq v \in \mathbb{C}^m$ which we call an *eigenvector*, where the symbol "$*$" denotes the usual matrix/vector multiplication.

Let $\mathcal{M}(m, n)$, $m \leq n$, be the linear space of all $(m \times n)$-matrices with complex entries. In what follows we will consider k-tuples of $(m \times n)$-matrices $B_1, \ldots, B_k$ which are linearly independent in $\mathcal{M}(m, n)$ and denote their linear span by $\mathcal{L} = \mathcal{L}(B_1, \ldots, B_k)$. Given a matrix pencil $\mathcal{P} = A + \mathcal{L}$, where $A \in \mathcal{M}(m, n)$, let $\mathcal{E}_\mathcal{P} \subset \mathcal{P}$ be its *eigenvalue locus*, i.e., the set of matrices in $\mathcal{P}$ whose rank is less than m. Elements of $\mathcal{E}_\mathcal{P}$ will be called *(generalized) eigenvalues*. Denote by $\mathcal{M}^1 \subset \mathcal{M}(m, n)$ the set of all $(m \times n)$-matrices with positive corank, i.e., whose rank is non-maximal. Its codimension equals $n - m + 1$ and its degree as an algebraic variety equals $\binom{n}{m-1}$, see [BV, Proposition 2.15]. Consider the natural left-right action of the group $GL_m \times GL_n$ on $\mathcal{M}(m, n)$, i.e., GL_m (respectively, GL_n) acts on $(m \times n)$-matrices by left (respectively, right) multiplication. This action on $\mathcal{M}(m, n)$ has finitely many orbits, each orbit being the set of all matrices of a given (co)rank, see, e.g., [AVG, Chap. I §2]. Note that by the well-known product formula for coranks the codimension of the set of matrices of rank r equals $(m - r)(n - r)$. Obviously, for any pencil $\mathcal{P}$ one has that the eigenvalue locus coincides with $\mathcal{E}_\mathcal{P} = \mathcal{M}^1 \cap \mathcal{P}$. Thus for a generic pencil $\mathcal{P}$ of dimension k the eigenvalue locus $\mathcal{E}_\mathcal{P}$ is a subvariety of $\mathcal{P}$ of codimension $n - m + 1$ if $k \geq n - m + 1$ and it is empty otherwise. The most interesting situation for applications occurs when $k = n - m + 1$, in which case $\mathcal{E}_\mathcal{P}$ is generically a finite set. From now on we assume that $k = n - m + 1$. Denoting as above by $\mathcal{L}$ the linear span of $B_1, \ldots, B_{n-m+1}$ we say that $\mathcal{L}$ is *transversal to* $\mathcal{M}^1$

if the intersection $\mathcal{L} \cap \mathcal{M}^1$ is finite and *non-transversal to $\mathcal{M}^1$* otherwise. Notice that due to the homogeneity of $\mathcal{M}^1$ any $(n-m+1)$-dimensional linear subspace $\mathcal{L}$ transversal to it intersects $\mathcal{M}^1$ only at 0 and that the multiplicity of this intersection at 0 equals $\binom{n}{m-1}$.

An important and most natural example of such a subspace $\mathcal{L}$ is motivated by the Heine-Stieltjes theory [He] and its higher order generalizations [BBS]. Denote by J_s, $s = 1, \ldots, n-m+1$, the $(m \times n)$-matrix whose entries are given by $a_{i,j} = 0$ if $i - j \neq s$ and 1 otherwise. We call J_s the *s-th unit matrix* or the *the s-th diagonal matrix*. Let us denote the linear span of $J_1, \ldots, J_{n-m+1}$ by $\mathfrak{L}$ and call $\mathfrak{L}$ the *standard diagonal subspace*. Note that $\mathfrak{L}$ is transversal to $\mathcal{M}^1$ since any matrix in $\mathfrak{L}$ different from 0 has full rank, as one can easily check.

We start with the following simple statement.

Lemma 1. *If $\mathcal{L} \subset \mathcal{M}(m,n)$ has dimension $(n-m+1)$ and is tranversal to $\mathcal{M}^1$ then for any matrix $A \in \mathcal{M}(m,n)$ the eigenvalue locus $\mathcal{E}_\mathcal{P}$ of the pencil $\mathcal{P} = A + \mathcal{L}$ consists of exactly $\binom{n}{m-1}$ points counted with multiplicities.*

Remark 1. Notice that since $\mathcal{M}^1$ is an incomplete intersection the same holds for the eigenvalue locus $\mathcal{E}_\mathcal{P}$ of a generic pencil $\mathcal{P} = A + \mathcal{L}$, i.e., in order to find $\mathcal{E}_\mathcal{P}$ for a given generic matrix A and a given generic subspace $\mathcal{L}$ one has to solve an overdetermined system of determinantal equations.

However, as was essentially discovered by Heine [He], the situation is different if one considers the standard diagonal subspace $\mathfrak{L}$ and any $A = (a_{i,j}) \in \mathcal{M}(m,n)$ which is upper-triangular – that is, such that $a_{i,j} = 0$ whenever $i > j$ – and has additionally distinct elements on the first main diagonal.

Theorem 1. *For any upper-triangular matrix $A = (a_{i,j}) \in \mathcal{M}(m,n)$ with all distinct entries $a_{i,i}$ on the first main diagonal the eigenvalue locus $\mathcal{E}_\mathcal{P}$ of the pencil $\mathcal{P} = A + \mathfrak{L}$, where $\mathfrak{L}$ is the standard diagonal subspace, is the union of m complete intersections enumerated by the first component of the eigenvalue.*

Remark 2. An explicit defining system of $(n-m)$ algebraic equations in $(n-m)$ variables for each such complete intersection is presented in the proof of Theorem 1, see §1 below.

Given $\mathcal{L}$ as above consider the natural projection map $\pi_\mathcal{L} : \mathcal{M}(m,n) \to \mathcal{L}^\perp$ along $\mathcal{L}$, where $\mathcal{L}^\perp = \mathcal{M}(m,n)/\mathcal{L}$. Noticing that $\dim \mathcal{M}^1 = \dim \mathcal{L}^\perp$ we define the *set of critical values* of $\pi_\mathcal{L}$ to be the set $\mathcal{C}_\mathcal{L}$ of all points in $\mathcal{M}^1$ where $\pi_\mathcal{L}$ is not a local diffeomorphism of $\mathcal{M}^1$ on its image $\pi_\mathcal{L}(\mathcal{M}^1)$. In other words, $\mathcal{C}_\mathcal{L}$ is the set of all points $p \in \mathcal{M}^1$ such that the sum of $\mathcal{L}$ and the tangent space to $\mathcal{M}^1$ at p does not coincide with the whole $\mathcal{M}(m,n)$. In particular, independently of $\mathcal{L}$ the critical value set $\mathcal{C}_\mathcal{L}$ always includes the set $\mathcal{M}^2$ of all $(m \times n)$-matrices with corank at least 2.

Recall that $\mathcal{M}^1 \subset \mathcal{M}(m,n)$ has the classical small resolution of singularities $\widetilde{\mathcal{M}^1} \subset \mathcal{M}(m,n) \times \mathbb{CP}^{m-1}$. Here $\widetilde{\mathcal{M}^1}$ consists of all pairs $(A, pker(A))$, where $A \in \mathcal{M}^1$ and $pker(A)$ is the projectivization of the left kernel of A. Using this construction one can parameterize a Zariski open subset of $\mathcal{M}^1$ as follows. Consider the product $P(m,n) = \mathcal{M}(m-1,n) \times \mathbb{C}^{m-1}$. Take the map $\nu : P(m,n) \to \mathcal{M}^1 \subset \mathcal{M}(m,n)$ sending a pair $(\mathcal{A}; k_1, \ldots, k_{m-1})$ to the matrix $A \in \mathcal{M}(m,n)$ obtained by appending to $\mathcal{A}$ the last row such that its sum with the linear combination with the coefficients $(k_1, \ldots, k_{m-1})$ of the respective rows of $\mathcal{A}$ vanishes.

The main result of this paper is a simple determinantal representation of $\mathcal{C}_\mathcal{L}$ in the above coordinates.

Theorem 2. *Let $\mathcal{L}$ be any $(n - m + 1)$-dimensional linear subspace in $\mathcal{M}(m, n)$ transversal to $\mathcal{M}^1$ and denote by $L_1, \ldots, L_{n-m+1}$ some basis of $\mathcal{L}$. Then in the coordinates of $P(m, n)$ the critical value set $\mathcal{C}_{\mathcal{L}}$ is given the determinantal equation*

$$\det \begin{pmatrix} \mathcal{A} \\ V_1 \\ \vdots \\ V_{n-m+1} \end{pmatrix} = 0. \tag{0.1}$$

Here $\mathcal{A}$ is a $(m-1, n)$-matrix with undetermined entries and V_j, $j = 1, \ldots, n-m+1$, are row vectors given by $V_j = \kappa * L_j$, where $\kappa = (k_1, \ldots, k_m)$.

Remark 3. If one expands equation (0.1) in the variables $(k_1, \ldots, k_{n-m+1})$ then the coefficient of each monomial in these variables is a linear combination of the maximal minors of $\mathcal{A}$ (i.e., the Plücker coordinates) with complex coefficients depending only on the choice of $\mathcal{L}$. Moreover, the above equation contains a lot of information of geometric nature.

Our next result shows that for the standard diagonal subspace $\mathfrak{L}$ the determinantal equation in Theorem 2 can be made quite a bit more explicit, which is particularly convenient from a computational viewpoint. We need first some additional notation. If $s \geq 1$ is an integer and $1 \leq r \leq s$ let $Q_{r,s}$ be the set of all strictly increasing sequences of r integers chosen from $1, \ldots, s$. Note in particular that $Q_{s,s}$ consists of a single sequence, namely $\{1, \ldots, s\}$. For $\alpha = (\alpha_1, \ldots, \alpha_r) \in Q_{r,s}$ set $\rho(\alpha) = \sum_{j=1}^{r} \alpha_j$. Given $A \in \mathcal{M}(m, n)$, $1 \leq k \leq m$, $1 \leq l \leq n$, $\alpha \in Q_{k,m}$ and $\beta \in Q_{l,n}$ denote by $A[\alpha|\beta] \in \mathcal{M}(k, l)$ the submatrix of A lying in rows α and columns β. Let $\mathcal{HP}(i, d)$ denote the complex space of all homogeneous polynomials in i variables of degree d and define the $d \times (i + d - 1)$ matrix

$$T_{i,d} = T_{i,d}(k_1, \ldots, k_i) = k_1 J_1 + \ldots + k_i J_i,$$

where J_j, $1 \leq j \leq i$, is as before the j-th diagonal $d \times (i+d-1)$ matrix and $k_1, \ldots, k_i$ are indeterminates. We will also need a result that may be of independent interest, namely the following lemma.

Lemma 2. *In the above notation, the $\binom{i+d-1}{d}$ polynomials in $k_1, \ldots, k_i$ given by the determinants*

$$\left| T_{i,d}[\alpha|\beta] \right|, \quad \alpha \in Q_{d,d}, \ \beta \in Q_{d,i+d-1},$$

build a basis of $\mathcal{HP}(i, d)$.

Remark 4. The usual determinant expansion formula provides an explicit expression (albeit tedious and not really needed for the present purposes) for the $\binom{i+d-1}{d} \times \binom{i+d-1}{d}$ matrix relating the standard monomial basis of $\mathcal{HP}(i, d)$ to the one constructed in Lemma 2.

Theorem 3. *Let $\mathcal{A} \in \mathcal{M}(m-1, n)$ be as in Theorem 2. The homogeneous defining polynomial of $\mathcal{C}_{\mathfrak{L}}$ with respect to the standard diagonal subspace $\mathfrak{L}$ is given by*

$$\sum_{\beta \in Q_{m-1,n}} (-1)^{\rho(\beta)} \left| \mathcal{A}\big[\{1, \ldots, m-1\}|\beta\big] \right| \cdot \left| T_{m,n-m+1}\big[\{1, \ldots, n-m+1\}|\{1, \ldots, n\} \backslash \beta\big] \right|.$$

Example 1. For $m = 2$ the homogeneous defining polynomial of $\mathcal{C}_{\mathfrak{L}}$ with respect to the standard diagonal subspace $\mathfrak{L}$ is given by

$$a_n k_1^n + a_{n-1} k_1^{n-1} k_2 + a_{n-2} k_1^{n-2} k_2^2 + \ldots + a_1 k_1 k_2^{n-1},$$

where $a_j = a_{1,j}$, $j = 1, \ldots, n$.

Example 2. For the standard diagonal subspace $\mathfrak{L}$ in the case of $\mathcal{M}(3,4)$ the homogeneous defining polynomial of $\mathcal{C}_{\mathfrak{L}}$ may be written as

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ k_1 & k_2 & k_3 & 0 \\ 0 & k_1 & k_2 & k_3 \end{vmatrix} = \Delta_{3,4}k_1^2 + \Delta_{1,4}k_2^2 + \Delta_{1,2}k_3^2 - \Delta_{2,4}k_1k_2$$

$$+ (\Delta_{2,3} - \Delta_{1,4})k_1k_3 - \Delta_{1,3}k_2k_3,$$

where $\Delta_{i,j}$ is the (2×2)-determinant of the upper part $\mathcal{A}$ including the i-th and j-th columns.

Remark 5. The multiplicity of an eigenvalue $A \in \mathcal{E}_{\mathcal{P}}$ can be expressed in terms of the dimension of the corresponding local algebra. More exactly, for an $(m \times n)$-matrix A we define the ideal I_A in the algebra $\mathbb{C}[[t_1, \ldots, t_k]]$ of formal power series as the ideal generated by all Plücker polynomials $\Delta_{i_1,\ldots,i_m}(A + \sum_{l=1}^m t_l B_l)$, where $\Delta_{i_1,\ldots,i_m}(X_{m \times n})$ is the determinant of the $(m \times m)$ submatrix of X formed by the columns with the indices $i_1, i_2, \ldots, i_m$. Now define the local algebra $\mathcal{A}_{loc}$ as the quotient algebra $A_{loc} = \mathbb{C}[[t_1, \ldots, t_k]]/I_A$. Then the multiplicity of the eigenvalue A in the pencil $A + \mathcal{L}$ equals $\dim_{\mathbb{C}} A_{loc}$.

The main result of this note (Theorem 2) gives a simple explicit determinantal formula for the critical value set $\mathcal{C}_{\mathcal{L}}$ (in coordinates on the resolution of singularities $P(m,n)$). Its inverse image $\pi_{\mathcal{L}}^{-1}(\mathcal{C}_{\mathcal{L}})$ is an important hypersurface consisting of all matrices in $\mathcal{M}(m,n)$ having a multiple eigenvalue. However, the problem of obtaining explicitly its defining polynomial in matrix entries seems to be quite delicate in general. As an illustration, let us show how this can be done in the simplest case of (2×3)-matrices.

Example 3 (Discriminant equation). For $m = 2, n = 3$ we will write the defining equation for the hypersurface $\pi_{\mathcal{L}}^{-1}(\mathcal{C}_{\mathcal{L}})$ of matrices with multiple eigenvalues in the space $\mathcal{M}(2,3)$ itself.

For any pair of positive integers $m < n$ consider the extended matrix space $\mathcal{M}(m,n) \times \mathbb{C}P^{m-1} \times \mathbb{C}^{n-m+1}$, where the m-tuple of homogeneous coordinates in $\mathbb{C}^m$ is denoted by $\kappa = (\kappa_1 : \cdots : \kappa_m)$ and the coordinates in $\mathbb{C}^{n-m+1}$ are denoted by $\lambda = (\lambda_1, \ldots, \lambda_{n-m+1})$.

Given a matrix $M \in \mathcal{M}(m,n)$ we will write a system of polynomial equations for $A + \sum_{i=1}^{n-m+1} \lambda_i J_i$,

$$\sum_{s=1}^m \kappa_s [A + \sum_{i=1}^{n-m+1} \lambda_i J_i]_{s*} = 0,$$

expressing the fact that λ is an eigenvalue of M while the κ_i's are the corresponding coefficients of a linear dependence between the rows of the matrix.

Using resultants we can get rid of the additional variables λ and κ. This elimination leads to the defining equation for the hypersurface in question.

Namely, consider a (2×3)-matrix $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$ and let as before $J_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ and $J_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. A generic element of the pencil $\mathcal{P}$ is thus given by

$$A(\lambda_1, \lambda_2) := A - \lambda_1 J_1 - \lambda_2 J_2 = \begin{pmatrix} a_{11} - \lambda_1 & a_{12} - \lambda_2 & a_{13} \\ a_{21} & a_{22} - \lambda_1 & a_{23} - \lambda_2 \end{pmatrix}.$$

For a generic matrix A the condition that the rank of $A(\lambda_1, \lambda_2)$ is less than 2 translates into two equations: the minor consisting of the second and third columns

vanishes, and the minor consisting of the first and third columns vanishes. These equations have the form

$$(a_{12} - \lambda_2)(a_{23} - \lambda_2) - a_{13}(a_{22} - \lambda_1) \quad = \quad 0 \qquad (0.2)$$
$$(a_{11} - \lambda_1)(a_{23} - \lambda_2) - a_{13}a_{21} \quad = \quad 0 \qquad (0.3)$$

Note that

$$\kappa := \kappa_1 = \frac{a_{23} - \lambda_2}{a_{13}}. \qquad (0.4)$$

Moreover, from the determinantal equation of Theorem 2 we obtain a third equation. Substituting expression (0.4) into the latter gives the equation

$$a_{13}^2 a_{11} - a_{13}^2 \lambda_1 + a_{23}a_{13}a_{12} - 3a_{23}a_{13}\lambda_2 + a_{13}a_{23}^2 - \lambda_2 a_{13}a_{12} + 2a_{13}\lambda_2^2 = 0. \quad (0.5)$$

Now equation (0.2) has bidegree $(1, 2)$ with respect to λ_1, λ_2. Analogously, (0.3) has bidegree $(1, 1)$ and (0.5) has bidegree $(1, 2)$ with respect to the same variables. Clearly, any solution s of the system of equations consisting of (0.2), (0.3) and (0.5) annihilates any polynomial in the ideal generated by these three equations. In particular, the following eight equations have s as a common solution: (0.2), (0.2) multiplied by λ_2, (0.2) multiplied by λ_2^2, (0.3), (0.3) multiplied by λ_2, (0.5), (0.5) multiplied by λ_2, (0.5) multiplied by λ_2^2. Therefore, the (8×8) determinant

$$D =$$

$$\det \begin{pmatrix} \Delta_{23} & a_{13} & -a_{12} - a_{23} & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \Delta_{23} & a_{13} & -a_{12} - a_{23} & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \Delta_{23} & a_{13} & -a_{12} - a_{23} & 1 \\ \Delta_{13} & -a_{23} & -a_{11} & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \Delta_{13} & -a_{23} & -a_{11} & 1 & 0 & 0 \\ \delta & -a_{13}^2 & \sigma & 0 & 2a_{13} & 0 & 0 & 0 \\ 0 & 0 & \delta & -a_{13}^2 & \sigma & 0 & 2a_{13} & 0 \\ 0 & 0 & 0 & 0 & \delta & -a_{13}^2 & \sigma & 2a_{13} \end{pmatrix}$$

vanishes when (0.2), (0.3) and (0.5) have a common root. Here we use the following notation: $\delta = a_{13}^2 a_{11} + a_{23}a_{13}a_{12} + a_{13}a_{23}^2$, $\Delta_{23} = a_{12}a_{23} - a_{13}a_{22}$, $\Delta_{13} = a_{11}a_{23} - a_{21}a_{13}$ and $\sigma = -a_{13}a_{12} - 3a_{13}a_{23}$.

This observation implies that the required defining polynomial for $\pi_{\mathcal{L}}^{-1}(\mathcal{C}_{\mathcal{L}})$ is the product of some (but not necessarily all) irreducible factors of the polynomial D. Factorizing D we obtain $D = a_{11}^6 D_0$, where

$$\begin{aligned} D_0 = & - 12a_{13}a_{22}^2 a_{11} + a_{22}^2 a_{12}^2 + 12a_{13}a_{22}a_{11}^2 + a_{11}^2 a_{23}^2 + 4a_{21}a_{12}^3 \\ & - 4a_{21}a_{23}^3 + a_{11}^2 a_{12}^2 + 12a_{12}a_{23}^2 a_{21} - 12a_{12}^2 a_{23}a_{21} - 2a_{12}a_{23}a_{22}^2 \\ & - 2a_{12}a_{23}a_{11}^2 - 2a_{22}a_{11}a_{23}^2 - 18a_{13}a_{22}a_{23}a_{21} - 2a_{22}a_{11}a_{12}^2 \\ & + 18a_{13}a_{22}a_{21}a_{12} + 18a_{11}a_{23}a_{13}a_{21} - 18a_{21}a_{13}a_{12}a_{11} \\ & + 4a_{13}a_{22}^3 - 27a_{21}^2 a_{13}^2 - 4a_{13}a_{11}^3 + 4a_{12}a_{23}a_{22}a_{11} + a_{22}^2 a_{23}^2 = 0 \end{aligned}$$

Note that D_0 is of second degree in the variable a_{13} and its discriminant (with respect to this variable) $W = 16(3a_{12}a_{21} - 3a_{21}a_{23} - 2a_{11}a_{22} + a_{11}^2 + a_{22}^2)^3$ is not a complete square. Thus, we conclude that D_0 is irreducible. Hence the variety given by $\{D = 0\}$ is the union of the variety given by $\{D_0 = 0\}$ and the hyperplane $\{a_{11} = 0\}$ taken with multiplicity 6.

Since the hyperplane $\{a_{11} = 0\}$ is obviously not contained in $\pi_{\mathcal{L}}^{-1}(\mathcal{C}_{\mathcal{L}}) \subset \mathcal{M}(2, 3)$ we obtain that $\pi_{\mathcal{L}}^{-1}(\mathcal{C}_{\mathcal{L}})$ is given by $\{D_0 = 0\}$.

1. PROOFS

Proof of Lemma 1. This follows almost directly from homogeneity of $\mathcal{M}^1$. Indeed, take any matrix $0 \neq A \in \mathcal{M}(m,n)$. Let $\tilde{l} \in \mathcal{L}$ be its eigenvalue, that is a matrix from $\mathcal{L}$ such that $A + \tilde{l}$ belongs to $\mathcal{M}^1$. Notice that for any $\epsilon \in (0,1]$ the matrix $\epsilon \tilde{l}$ is the eigenvalue of the matrix ϵA. Considering the family of matrices ϵA with $\epsilon \in [0,1]$ we conclude that the total multiplicity of eigenvalues of the pencil $A + \mathcal{L}$ coincides with that of the linear pencil $\mathcal{L}$ if the latter multiplicity is finite, which gives the required statement. $\square$

Proof of Theorem 1. To get the defining system of algebraic equations for $\mathcal{E}_{\mathcal{P}}$ under the assumptions of Theorem 1 we proceed exactly as in [He]. For a given upper-triangular matrix $A \in \mathcal{M}(m,n)$ with distinct entries on the main diagonal we want to find all $(n-m+1)$-tuples $(\lambda_1, \ldots, \lambda_{n-m+1})$ such that the matrix $A + \lambda_1 J_1 + \lambda_2 J_2 + \ldots + \lambda_{n-m+1} J_{n-m+1}$ has positive corank. Since A is upper-triangular with distinct $a_{i,i}$ then in order to get a positive corank it is necessary to require $\lambda_1 + a_{i,i} = 0$ for some $i = 1, \ldots, m$. The next observation is that under the above assumptions on A for any given $i = 1, \ldots, m$ the total number of eigenvalues with $\lambda_1 + a_{i,i} = 0$ equals $\binom{n-i}{m-i}$ which gives the following count of the eigenvalues of A noticed already by Heine: $\binom{n}{m-1} = \binom{n-1}{m-1} + \binom{n-2}{m-2} + \ldots + \binom{n-m}{0}$. Indeed, if $\lambda_1 + a_{i,i} = 0$ then $\lambda_1 + a_{j,j} \neq 0$ for all $j \neq i$ and, in particular due to the assumptions on A the first $i-1$ rows of $A - a_{i,i} J_1 + \lambda_2 J_2 + \ldots + \lambda_{n-m+1} J_{n-m+1}$ are linearly independent for all values of $\lambda_2, \ldots, \lambda_{n-m+1}$. On the other hand, the remaining rows $i, i+1, \ldots, m$ can become linearly dependent under an appropriate choice of $\lambda_2, \ldots, \lambda_{n-m+1}$. Since the matrix $A - a_{1,1} J_1$ is upper-triangular with the (i,i)-th entry vanishing the condition that $A - a_{i,i} J_1 + \lambda_2 J_2 + \ldots + \lambda_{n-m+1} J_{n-m+1}$ has positive corank is equivalent to the condition that the matrix obtained by removing its first i rows and $i-1$ columns has positive corank. By Lemma 1 the total number of eigenvalues of the matrix of the size $(m-i+1) \times (n-i)$ equals $\binom{n-i}{m-i}$. Let us now for any given $i = 1, \ldots, m$ derive a system of algebraic equations in the variables $\lambda_2, \ldots, \lambda_{n-m+1}$ whose solutions are exactly all the eigenvalues of A with $\lambda_1 + a_{i,i} = 0$. We will concentrate on the case $i = 1$ since all other cases are covered in exactly the same way by working with a smaller matrix obtained from A by removing the first $(i-1)$ rows and $(i-1)$ columns. Using $(k_1, \ldots, k_m)$ for the coordinates of the left kernel and $\lambda_1, \lambda_2, \ldots, \lambda_{n-m+1}$ for the eigenvalues we get the following system of equations

$$
\begin{cases}
0 = k_1(a_{1,1} + \lambda_1) \\
0 = k_1(a_{1,2}\lambda_2) + k_2(a_{2,2} + \lambda_1) \\
\quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
0 = k_1(a_{1,m} + \lambda_m) + k_2(a_{2,m} + \lambda_{m-1}) + \ldots + k_m(a_{m,m} + \lambda_1) \\
0 = k_1(a_{1,m+1} + \lambda_{m+1}) + k_2(a_{2,m+1} + \lambda_m) + \ldots + k_m(a_{m,m+1} + \lambda_2) \\
\quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
0 = k_1(a_{1,m+1} + \lambda_{m+1}) + k_2(a_{2,m+1} + \lambda_m) + \ldots + k_m(a_{m,m+1} + \lambda_2) \\
0 = k_1(a_{1,n} + \lambda_n) + k_2(a_{2,n} + \lambda_{n-1}) + \ldots + k_m(a_{n,n} + \lambda_{n-m+1})
\end{cases}
$$

expressing the existence of a nontrivial left kernel of $A + \lambda_1 J_1 + \lambda_2 J_2 + \ldots + \lambda_{n-m+1} J_{n-m+1}$. (To simplify notations we assume here that $\lambda_j = 0$ for $j > n - m + 1$.) In order to get the required system of equations in $\lambda_1, \ldots, \lambda_{n-m+1}$ we have to eliminate from the above system the variables $k_1, \ldots, k_m$. Notice that under our assumptions on A the possible corank of $A + \lambda_1 J_1 + \lambda_2 J_2 + \ldots + \lambda_{n-m+1} J_{n-m+1}$ can be at most 1 and in the case of corank 1 the linear dependence must necessarily include the first row, i.e., $k_1 = 1$. Note also that the first m equations are triangular with respect to $k_1, \ldots, k_m$, which together with our assumptions on A allows us to

successfully eliminate them. Namely, from the first equation we get $\lambda_1 = -a_{1,1}$ and $k_1 = 1$. Then for any $i = 2, \ldots, m$ we solve the i-th equation with respect to k_i and get

$$k_i = \frac{1}{a_{1,1} - a_{i,i}} \left(k_1(a_{1,i} + \lambda_i) + k_2(a_{2,i} + \lambda_{i-1}) + \ldots + k_{i-1}(a_{i-1,i} + \lambda_2) \right).$$

With the initial value $k_1 = 1$ and taking into account that the only possible denominators occurring in the above expressions for k_i are $a_{1,1} - a_{i,i}$ we recurrently find all k_i, $i = 1, \ldots, m$ as the functions of the matrix entries and λ's. Substituting these found expressions in the remaining $n - m$ equations we get the required system of algebraic equations to determine $\lambda_2, \ldots, \lambda_{n-m+1}$. (Notice that $\lambda_1 = -a_{1,1}$ was already obtained from the first equation.) $\qquad \square$

Example 4. Any matrix $A \in \mathcal{M}(2, 4)$ has four eigenvalues (counted with multiplicities) with respect to the standard diagonal subspace $\mathfrak{L}$. If A is upper-triangular with distinct elements on the first main diagonal then these eigenvalues split into two groups depending on the value of λ_1. Namely, there are 3 eigenvalues for which $\lambda_1 = -a_{1,1}$ and 1 eigenvalue for $\lambda_1 = -a_{2,2}$. For $\lambda_1 = -a_{1,1}$ the above system (before elimination) has the form:

$$\begin{cases} 0 = k_1(a_{1,1} + \lambda_1) \\ 0 = k_1(a_{1,2} + \lambda_2) + k_2(a_{2,2} + \lambda_1) \\ 0 = k_1(a_{1,3} + \lambda_3) + k_2(a_{2,3} + \lambda_2) \\ 0 = k_1 a_{1,4} + k_2(a_{2,4} + \lambda_3). \end{cases}$$

From the first equation we get $k_1 = 1$ and $\lambda_1 = -a_{1,1}$. From the second equation we get $k_2 = \frac{a_{1,2} + \lambda_2}{a_{1,1} - a_{2,2}}$. Substituting in the remaining two equations we get the next system to determine λ_2 and λ_3:

$$\begin{cases} (\lambda_2 + a_{1,2})(\lambda_2 + a_{2,3}) + (a_{1,1} - a_{2,2})(\lambda_3 + a_{1,3}) = 0 \\ (\lambda_2 + a_{1,2})(\lambda_3 + a_{2,4}) + (a_{1,1} - a_{2,2})a_{1,4} = 0. \end{cases}$$

In the case $\lambda_1 + a_{2,2} = 0$ one gets a very simple linear system:

$$k_2(a_{2,2} + \lambda_1) = k_2(a_{2,3} + \lambda_2) = k_2(a_{2,4} + \lambda_3) = 0$$

which gives $k_2 = 1$, $\lambda_1 = -a_{2,2}$, $\lambda_2 = -a_{2,3}$, $\lambda_3 = -a_{2,4}$.

Proof of Theorem 2. As we already mentioned in the introduction the set $\mathcal{C}_\mathcal{L}$ can be determined as the set of all matrices $M \in \mathcal{M}^1$ such that the sum of the tangent space to $\mathcal{M}^1$ at M and the linear space $\mathcal{L}$ does not coincide with the whole $\mathcal{M}(m, n)$. Let us describe a basis of the tangent space to $\mathcal{M}^1$ at a sufficiently generic matrix M. Since $GL_m \times GL_n$ acts on $\mathcal{M}(m, n)$ with finitely many orbits the tangent space to the $GL_m \times GL_n$-orbit of M under this action coincides with the tangent space to $\mathcal{M}^1$ at M. Note that $GL_m \times GL_n$ acts on $\mathcal{M}(m, n)$ by elementary row and column operations. Thus, if we take for example the affine chart in which the determinant formed by the first $(m - 1)$ rows and columns is non-vanishing then the tangent space to $\mathcal{M}^1$ at any matrix M belonging to this chart is generated by the following two groups of operations: (i) add to each column of M one of its first $m - 1$ columns and (ii) add to the last row of M one of its other rows. One has therefore a total of $n(m - 1) + (m - 1) = (n + 1)(m - 1) = \dim \mathcal{M}^1$ generators. Taking the wedge of these generators with the chosen basis of $\mathcal{L}$ and representing an $(m \times n)$-matrix as a mn-vector by patching together its rows we obtain the following $(mn \times mn)$-matrix that has a block structure of an $(m \times m)$-matrix with $(n \times n)$-blocks of the form given below:

$$\mathfrak{D} = \begin{pmatrix}
a_{1,1}I_n & a_{2,1}I_n & \cdots & a_{m-1,1}I_n & -\sum_{j=1}^{m-1} k_j a_{j,1} \cdot I_n \\
a_{1,2}I_n & a_{2,2}I_n & \cdots & a_{m-1,2}I_n & -\sum_{j=1}^{m-1} k_j a_{j,2} \cdot I_n \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
a_{1,m-1}I_n & a_{2,m-1}I_n & \cdots & a_{m-1,m-1}I_n & -\sum_{j=1}^{m-1} k_j a_{j,m-1} \cdot I_n \\
0_{m-1,n} & 0_{m-1,n} & \cdots & 0_{m-1,n} & \mathcal{A}_{m-1,n} \\
L_{1,1} & L_{2,1} & \cdots & L_{m-1,1} & L_{m,1} \\
L_{1,2} & L_{2,2} & \cdots & L_{m-1,2} & L_{m,2} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
L_{1,n-m+1} & L_{2,n-m+1} & \cdots & L_{m-1,n-m+1} & L_{m,n-m+1}
\end{pmatrix}.$$

Here $\mathcal{A} = \mathcal{A}_{m-1,n} = (a_{i,j})$, $i = 1, \ldots, m - 1$, $j = 1, \ldots, n$, I_n is the identity $(n \times n)$-matrix, $0_{m-1,n}$ is the $((m - 1) \times n)$-matrix with all vanishing entries, and, finally, $L_{i,j}$ is the i-th row of the matrix L_j, see Theorem 2. Notice that the determinant $\det(a_{i,j}I_n)$, $i = 1, \ldots, m - 1$, $j = 1, \ldots, m - 1$, of the upper-left block of $\mathfrak{D}$ equals Δ^{m-1}, where $\Delta = \det(a_{i,j})$, $i = 1, \ldots, m - 1$, $j = 1, \ldots, m - 1$, is the leftmost principal minor of $A_{m-1,n}$. By the above assumption the matrix $\mathcal{A}$ lies in the affine chart where $\Delta \neq 0$. Finally, we clear the low-left block $(L_{i,j})$, $i = 1, \ldots, m - 1$, $j = 1, \ldots, n - m + 1$, of $\mathfrak{D}$ by "killing" all its elements through row operations using the above upper-left block (which is a square and non-degenerate $((m - 1)n \times (m - 1)n)$-matrix) to obtain the low-right block coinciding exactly with the matrix in formula (0.1). Thus the determinant of the whole matrix $\mathfrak{D}$ equals the product between Δ^{m-1} and the determinant from Theorem 2. Since in the considered chart one has $\Delta \neq 0$ the result follows. $\qquad\square$

Proof of Lemma 2. Set $t = i + d$, so that $t \geq 2$. Since

$$\dim \mathcal{HP}(i, d) = \binom{i + d - 1}{d}$$

we have to show that the polynomials constructed in the lemma are linearly independent, which we prove this by induction on t. Note that this is trivially true for $t = 2$. Assume that it holds for some $t \geq 2$ and let i, d be such that $i + d = t + 1$. Suppose that $c_{\alpha\beta} \in \mathbb{C}$ are such that

$$\sum_{\substack{\alpha \in Q_{d,d} \\ \beta \in Q_{d,i+d-1}}} c_{\alpha\beta} \Big| T_{i,d}(k_1, \ldots, k_i)[\alpha|\beta] \Big| = 0.$$

Clearly, this may be rewritten as

$$\sum_{\substack{\alpha \in Q_{d,d} \\ \beta \in Q_{d,i+d-2}}} c_{\alpha\beta} \Big| T_{i,d}(k_1, \ldots, k_i)[\alpha|\beta] \Big|$$

$$+ k_i \cdot \sum_{\substack{\alpha \in Q_{d-1,d-1} \\ \beta \in Q_{d-1,i+d-2}}} c_{\alpha\beta} \Big| T_{i,d-1}(k_1, \ldots, k_i)[\alpha|\beta] \Big| = 0. \quad (1.1)$$

In particular, setting $k_i = 0$ we get

$$\sum_{\substack{\alpha \in Q_{d,d} \\ \beta \in Q_{d,i+d-2}}} c_{\alpha\beta} \Big| T_{i-1,d}(k_1, \ldots, k_{i-1})[\alpha|\beta] \Big| = 0$$

hence $c_{\alpha\beta} = 0$, $\alpha \in Q_{d,d}$, $\beta \in Q_{d,i+d-2}$, by the induction assumption since $(i-1)+d = t$. Together with (1.1) this implies that

$$\sum_{\substack{\alpha \in Q_{d-1,d-1} \\ \beta \in Q_{d-1,i+d-2}}} c_{\alpha\beta} \Big| T_{i,d-1}(k_1,\ldots,k_i)[\alpha|\beta] \Big| = 0,$$

which in turn yields $c_{\alpha\beta} = 0$, $\alpha \in Q_{d-1,d-1}$, $\beta \in Q_{d-1,i+d-2}$, again by the induction hypothesis since $i + (d-1) = t$. We conclude that $c_{\alpha\beta} = 0$ for all $\alpha \in Q_{d,d}$ and $\beta \in Q_{d,i+d-1}$, which proves the desired statement hence also the lemma. $\square$

Proof of Theorem 3. We will use the setting and notation of Lemma 2 with $i = m$ and $d = n - m + 1$. Fix the sequence $\alpha = \{1,\ldots,m-1\} \in Q_{m-1,n}$. Now consider the left-hand side of the determinantal equation in Theorem 2 in the case when $\mathfrak{L}$ is the standard diagonal subspace and $L_j = J_j$, $1 \leq j \leq n - m + 1$. In view of the generalized Laplace expansion theorem, see, e.g., [MM, §2.4.11], when expanding it by the rows α this left-hand side becomes

$$(-1)^{\frac{m(m-1)}{2}} \sum_{\beta \in Q_{m-1,n}} (-1)^{\rho(\beta)} \Big| \mathcal{A}\big[\{1,\ldots,m-1\}|\beta\big] \Big| \times$$

$$\Big| T_{m,n-m+1}(k_1,\ldots,k_m)\big[\{1,\ldots,n-m+1\}|\{1,\ldots,n\}\setminus\beta\big] \Big|,$$

which proves the theorem. $\square$

2. Remarks and open questions

A. By analogy with the above case, for a given triple n, m, r one can also consider $(m-r)(n-r)$-dimensional pencils of matrices in $\mathcal{M}(m,n)$ and study their intersections with the subvariety $\mathcal{M}^r$ of all matrices of corank at least r. In particular, a natural question is to find an analog of Theorem 2 in this situation.

B. It would be interesting to determine the equation for $\pi_{\mathcal{L}}^{-1}(\mathcal{C}_{\mathcal{L}})$ in general, see Example 3 in the Introduction. Another important direction is to determine the local multiplicity of a given eigenvalue in terms of the defining polynomial for $\mathcal{C}_{\mathcal{L}}$. Is there any analog of the Jordan normal form allowing to determine the multiplicity of a given eigenvalue?

C. Notice that the left-right action of $GL_m \times GL_n$ extends from the space $\mathcal{M}(m,n)$ to every space of (in)complete flags in $\mathcal{M}(m,n)$. For simple dimensional reasons, in most cases this action cannot have finitely many orbits.

Problem 1. On which spaces of (in)complete flags the above left-right action of $GL_m \times GL_n$ has finitely many orbits?

References

[AVG] V. Arnold, A. Varchenko, S. Gusein-Zade, Singularities of differentiable maps. Vol. I. The classification of critical points, caustics and wave fronts. Monogr. Math. **82**, Birkhäuser Boston, Inc., Boston, MA, 1985.

[BBS] J. Borcea, P. Brändén, B. Shapiro, *Algebraic and geometric aspects of Heine-Stieltjes theory*, in preparation.

[BEGM] G. Boutry, M. Elad, G. Golub, P. Milanfar, *The generalized eigenvalue problem for nonsquare pencils using a minimal perturbation approach*, SIAM J. Matrix Anal. Appl. **27** (2005), 582–601.

[BV] W. Bruns, U. Vetter, Determinantal rings. Lect. Notes Math. **1327**, Springer-Verlag, Berlin, 1988.

[CG] D. Chu, G. Golub, *On a generalized eigenvalue problem for nonsquare pencils.* SIAM J. Matrix Anal. Appl. **28** (2006), 770–787.

[He] E. Heine, Handbuch der Kugelfunctionen. Vol.1, pp. 472–479, Berlin: G. Reimer Verlag, 1878.

[MM]   M. Marcus, H. Minc, A survey of matrix theory and matrix inequalities. Allyn and
        Bacon, Inc., Boston, MA, 1964.
[TW]   L. Trefethen, T. Wright, *Pseudospectra of rectangular matrices*, IMA J. Numer. Anal.
        22 (2002), 501–519.
[Vol]   H. Volkmer, Multiparameter eigenvalue problems and expansion theorems. Lect. Notes.
        Math. **1356**, Springer-Verlag, 1988.

DEPARTMENT OF MATHEMATICS, STOCKHOLM UNIVERSITY, SE-106 91 STOCKHOLM, SWEDEN
E-mail address: julius@math.su.se

DEPARTMENT OF MATHEMATICS, STOCKHOLM UNIVERSITY, SE-106 91 STOCKHOLM, SWEDEN
E-mail address: shapiro@math.su.se

DEPARTMENT OF MATHEMATICS, MICHIGAN STATE UNIVERSITY, EAST LANSING, MI 48824-
1027, USA
E-mail address: mshapiromath.msu.edu